



The following document describes the procedures implemented by **Purdue University Libraries Archives & Special Collections** to prepare files for loading into *Content DM*. Before items can be loaded for public (or restricted) use, descriptions must be assigned to each image, and load directories and files must be created. This is the final check of the digital files; precise and consistent application of the following procedures is necessary for accurate and efficient processing of digital archives.

Purdue University Libraries Archives & Special Collections
File Load Preparation Procedures
for digital archives



<http://www.lib.purdue.edu/spcol>

© 2006 Purdue University, West Lafayette, Indiana

Last Revised: July 20, 2006

Prepared for in-house use by Liz Erlewine under the supervision of Carl Snow, Digital Initiatives Librarian, and Sammie Morris, Archivist.

FINDING FILES

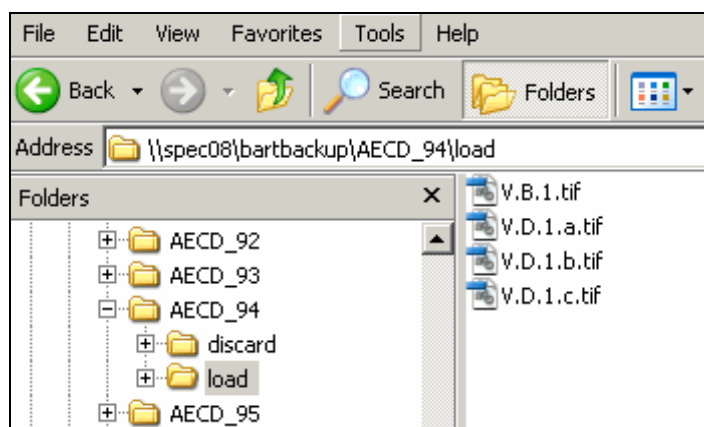
Files for loading will be requested by the Digital Initiatives Librarian and can be found on **SPEC08 Bartbackup**. These files have already been scanned and have been through Quality Control (*see Quality Control Procedures*).

Files will be requested by **box number, collection name, or CD number**. Look on Bartbackup to find the requested files.

LOAD FOLDER

Once you have located files that need to be loaded, create a new folder within it and name the new folder **Load**.

Move all files that need to be loaded into the new **Load** folder.



LOAD FILE

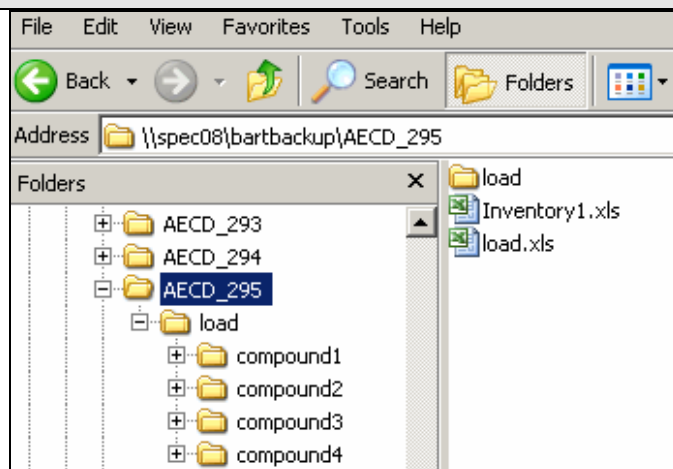
Images within the **Load** folder should be listed in an *Excel* spreadsheet titled **load.xls**

You can copy Loadsource.xls from SPEC08 Bartbackup and rename it load.xls.

Loadsource.xls is a blank version of a **load.xls** file with formulas for W-cm, H-cm, Source, Orientation, and Object ID.

LOAD FILE LOCATION

Load.xls should be located in the directory directly **above the Load folder**.



LOAD.XLS CONTENTS

Load.xls should contain the information listed below.

*Most of the information can be copied from the **Inventory.xls** file. Exceptions are noted below.*

Old ID:	<i>If applicable</i>
ID:	ID number (example: aepb1f12i17)
Title:	Leave as Title . Titles are determined by the Meta Data Librarian.
Description:	Varies by collection. You may or may not be instructed to add data to this field. Check with your supervisor, and <i>see Item Descriptions instructional document.</i>
Date Scanned:	Must be text format, MM/DD/YYYY .
Width:	In inches. (<i>required for images, not text</i>)
Height:	In inches. (<i>required for images, not text</i>)
W-cm:	You do not have to enter a value here. Use loadsource.xls for formula to convert the Width in inches to width in centimeters.
H-cm:	You do not have to enter a value here. Use loadsource.xls for formula to convert the Width in inches to width in centimeters.
Source:	Displays size in centimeters, H x W. Use loadsource.xls for formula.
Orientation:	Portrait or Landscape. Use loadsource.xls for formula.
Resolution:	Image resolution, pixels per inch: 600 ppi or 300 ppi.
Burned To:	Name of the CD where the image is stored (if already burned to disc).
Object ID:	File name of the image. Matches ID number (unless item is a compound document). Should always be TIF file and end with the extension .tif.

COMPOUND DOCUMENTS

Items with **multiple pages** or parts are considered **compound documents** and must be prepared a bit differently.

Compound documents must be grouped, renamed, and stored in separate **compound** folders, numbered sequentially.

Note: If you come across a document broken into chapters, check with a supervisor before preparing the file for loading. The document may be considered a “monograph” item and may require different load preparation.

COMPOUND FOLDER

If you come across an item that consists of two or more images for different pages or parts (a compound document), create a new folder within the **Load** folder and name it **Compound1**.

Move the files that make up the compound document into the **Compound1** folder.

Additional compound documents should be treated the same way:

Each set of compound images should be located in a compound folder, numbered sequentially: compound1, compound2, compound3, etc.

RENAME FILES

Rename the files in the **Compound1** folder to describe each image, following the guidelines below.

Filenames consist of a number and name.

The first three numbers should be sequential for the item, indicating the order in which the images should be viewed. That number should be followed by an underscore (_) and then the descriptive name, such as page1 or page2.

Examples: **000_page1.tif**
 001_page2.tif
 002_page3.tif

If “page” does not appropriately describe your item and image, you may substitute alternative titles.

Possible alternative names:

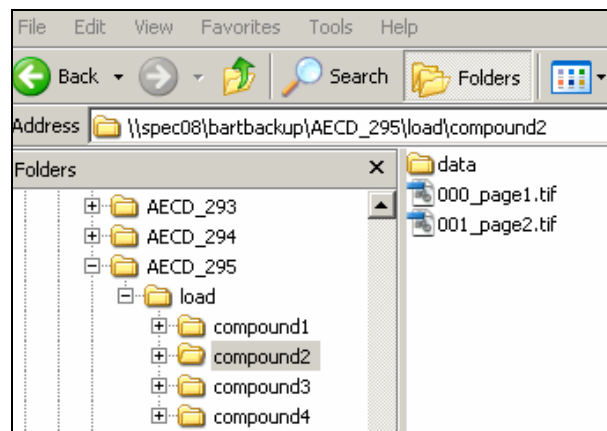
- Page
- View
- Front
- Back
- Inside
- Outside
- TOC
- Cover_Letter
- Map

When using alternate file names, you must still include the preceding sequential number and underscore. Note: the end user will not see this number, they will only see the descriptive name you give each file.

DATA FOLDER

Create a new folder within the compound folder and name it **Data**. (example: Bartbackup\AECD_295\Load\Compound2\Data)

Example of folder structure:



LOAD.XLS

Compound documents also require their own **load.xls** file.

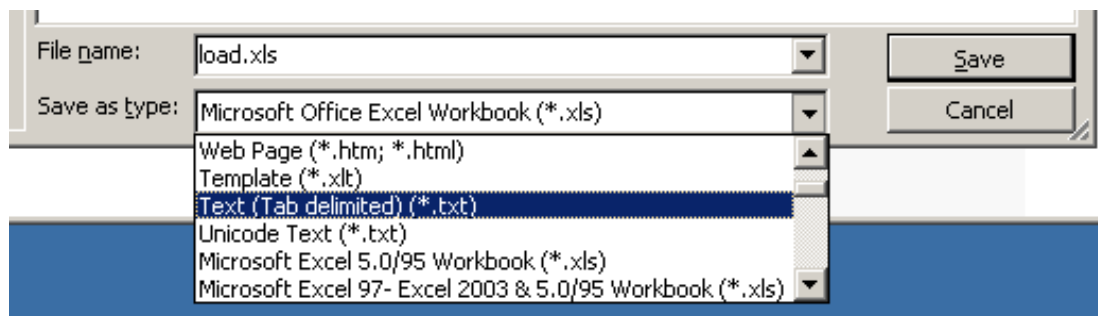
1. Copy the **load.xls** file from the main folder (example: Bartbackup\Amelia Earhart at Purdue) and paste it into the new **Data** folder:
(Bartbackup\AECD_123\Load\Compound1\Data).
2. Edit the **load.xls** file so that it only includes information for the compound documents located in the **Compound1** folder.
3. Change the **Object IDs** to match the new file names (*example: 000_page1.tif, 001_page2.tif*).
4. Insert a new row above the first item entry. Duplicate the information for the first item and paste it into the row you just added.
5. Save the file.

LOAD.TXT

In addition to the *Excel* spreadsheet, save the load file as a **Text (tab delimited)** file.

To do this: When in *Excel*, Click on **File** and then **Save**.
Select **Save As**.

For **File Type**, select **Text (tab delimited)**.



The **load.txt** file should be saved in the **Data** folder along with **load.xls**.

ADDITIONAL COMPOUND OBJECTS

Repeat these steps for all compound objects, creating sequential folders: **Compound1**, **Compound2**, **Compound3**, etc...

*Remember: Each **Compound** folder should contain a **Data** folder with a **load.xls** and **load.txt** file, and each file in the **Compound** folder should follow the naming conventions described above.*

CDM LOAD LIST

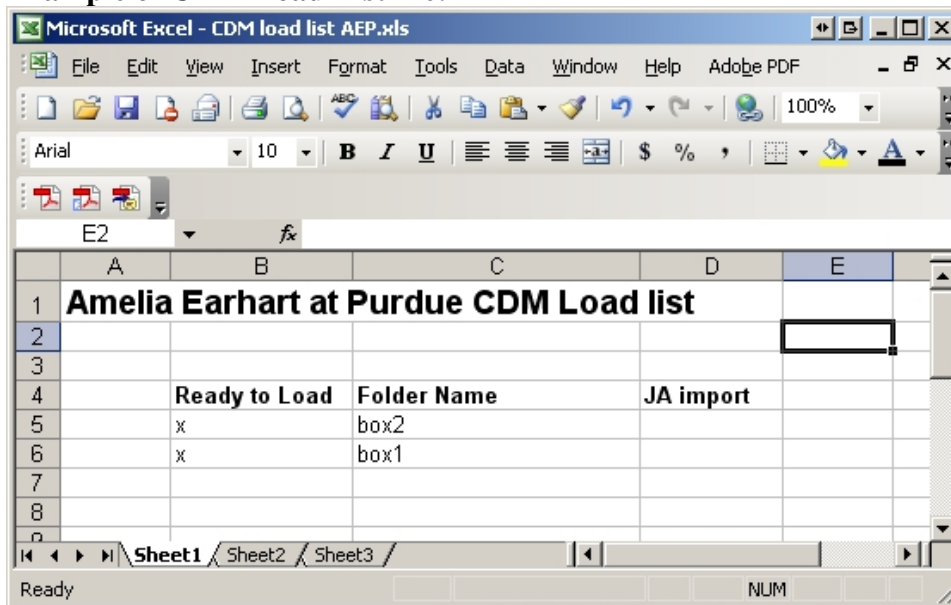
Once the images are ready for loading, you must update (or create) a CDM load list file for the collection. This file is located on **SPEC08 Bartbackup**.

The CDM load list is an *Excel* spreadsheet, listing the folders that are ready for loading.

Find the **CDM Load List** for the collection you are working on, and add the folder to the list.

Put an X next to the completed folder, indicating the folder has been prepared for loading.

Example of CDM Load List file:



The screenshot shows a Microsoft Excel window titled "Microsoft Excel - CDM load list AEP.xls". The spreadsheet has the following data:

	A	B	C	D	E
1	Amelia Earhart at Purdue CDM Load list				
2					
3					
4		Ready to Load	Folder Name	JA import	
5		x	box2		
6		x	box1		
7					
8					
9					
10					

LOAD PREPARATION COMPLETE

You are now finished preparing the digital archives for loading into *Content DM*.